DOCUMENT RESUME

ED 093 950 TM 003 768

AUTHOR Huberty, Carl J.
TITLE Use of Some Discriminant Analysis Computer Programs.
PUB DATE Apr 74
NOTE 22p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, Illinois, April, 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS *Classification; *Computer Programs; Correlation; Data Analysis; *Discriminant Analysis; *Educational Research; Sampling; Statistical Analysis

ABSTRACT

There are two distinct but related purposes for carrying out a "discriminant analysis": (1) discrimination, and (2) classification. The primary objective of this paper was to review the outputs of selected computer programs often used to carry out a "discriminant analysis" with respect to these two purposes. Information provided by the programs on requisite date conditions for each type of analysis is discussed. The evidence indicates that to say one has carried out a "discriminant analysis" when using any of the selected programs would be misleading. The information obtained from any of the program is quite inadequate for either of the two purposes mentioned above. (Author)

ED 093950

TM 003 768

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

BEST COPY AVAILABLE

USE OF SOME "DISCRIMINANT ANALYSIS" COMPUTER PROGRAMS

Carl J Huberty

University of Georgia

Session No. 28.02

BEST COPY AVAILABLE

Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, April, 1974.

ABSTRACT

The primary objective of this paper is to review the outputs of selected computer programs often used to carry out a "discriminant analysis" with respect to two purposes of such analysis: 1) discrimination, and 2) classification. The programs selected are the three BMD programs. Information provided by the programs in terms of requisite data conditions for each type of analysis is discussed. It is concluded that to say one has carried out a "discriminant analysis" when using any of the selected programs would be misleading, indeed. The information yielded directly by any of the programs is quite inadequate for either of the two purposes mentioned above. The obtaining of supplemental statistics is indicated.

Use of Some "Discriminant Analysis" Computer Programs

Introduction

Multivariate statistical theory is by no means new. However, applications of many apsects of the theory in educational research have only become fairly commonplace in the past decade or so. Interest in the complicated (in the sense of calculations, at least) multivariate procedures has certainly been enhanced by the adaptation of high speed computers to problems of data analysis. Except for very "small" sets of data, there has been almost a total reliance on computers by educational researchers to carry out the necessary calculations. In some cases of multivariate data analysis, problems have arisen out of the widespread use of computer programs. It must be noted that the problems are usually not inherent in the programs themselves, but rather in how they are used; albeit insufficient program documentation sometimes causes difficulty in use. Ofter times, lack of statistical training and/or experience in data analysis contribute to the misuse of computer programs, including misinterpretation of computer output. Problems with, and misuse of, computer programs have often appeared in two classes of multivariate methods: factor analysis and discriminant analysis. The concern in this paper is with the latter of these two general and often confusing domains of study.

Discriminant Analysis

The term "discriminant analysis" has come to mean different things to different people. The original proposed use of the "linear discriminant function" was to classify an object into one of two groups to which it must belong (Fisher, 1936). This classification is made using measures on a number of (intercorrelated) variables for each object involved. Even with more than two criterion groups, 'discriminant analysis" in educational applications has, in the past, most generally implied some type of classification or assignment of individuals. However, recently the term has taken on extended meaning; that is, the term may imply data analysis techniques other than mere classification. Suppose we are given the existence of g well-defined populations and a sample (or group) of individuals from each population with p measures for each individual. Methods used to analyze such data may be dictated by two purposes of the analysis: (1) to study group separation in terms of variable contribution and in terms of dimensions of separation (discrimination), and (2) to set up a rule, based on the p-variate data, which will enable us to assign some new individual to the correct population when it is not known from which of the g populations he emanates (classification). It may be added that two other purposes might be considered: (3) to determine if the g populations are statistically significantly separated (separation), and (4) to estimate distances between pairs of populations (estimation). It may be argued that separation-- a la multivariate analysis of variance (MANOVA) -- is necessarily considered prior to discrimination. See Huberty (1974) for a more complete discussion of these four aspects of discriminant analysis.

Studies designed with either of the first two purposes in view are scattered throughout the educational research literature. Discrimination analyses have recently been employed by Goldman and Warren (1973), Nicholson (1973), Whellams (1973), Bausell and Magoon (1972), and Rock, Baird, and Linn (1972). Classification was the primary analysis used by Keenen and Holmes (1970), Stahmann (1969), and Ghastian (1969). It must be recognized that not all studies which might be included in the latter category employ a classification analysis for the purpose mentioned in the previous paragraph. Rather, the individuals being classified are those whose measures were used in determining the classification rule applied. More will be said on this later.

Requisite Conditions for Discriminant Analysis

A wealth of research has been reported where the effects of failing to meet requisite conditions for univariate parametric statistical methods have been studied. The conditions usually considered in these studies are those of population normality and homogeneity of variance. In the univariate case very substantial departures from normality and/or homogeneity do not seem to affect many tests; at least in some senses. It is not at all clear that this holds in multivariate tests; relatively little empirical research has been done in this area.

A "discriminant analysis" in the sense of discrimination and classification problems may be carried out without directly incorporating significance tests. [However, some methodologists might contend that either of these two problems ought only be considered after a simple MANOVA yields significance.] The conditions of p-variate normality and equality of the g population (pxp) covariance matrices are often assumed to be met

(in some respects needlessly) in many discriminant analyses. Of course no such assumptions need be made in arriving at the sets of discriminant function[1] coefficients through the usual eigenanalysis. The sets of coefficients are the eigenvectors associated with the eigenvalues of the matrix product $W^{-1}B$, where W and B are the (pxp) pooled within-groups and between-groups deviation score cross-products matrices, respectively. It might be argued that such pooling only makes sense when the population covariance matrices are identical; it is noted that Porebski (1966, p. 228) debates the need for carrying out a preliminary test for identical population covariance matrices. In discrimination, the p-variate normality condition is only needed if one desires, or feels compelled to, test the discriminant functions for significance.

In classification applications p-variate normality is not a requirement; it is only necessary that the population density functions be known (Melton, 1963). However, most of the distribution-based formulations developed by mathematical statisticians for classification purposes are built on multivariate normal densities. [Limited developments have been made which are distribution-free in nature (see Kendall, 1966).] The inequality of the covariance matrices presents no problem in multivariate classification. In fact, differences in variances and covariances can be very useful in improving classification accuracy. This is particularly true when there is considerable overlap among the groups. An added assumption that is often made in a classification analysis in educational research is that costs of misclassifying individuals associated with each of the g groups are identical. The situation of unequal costs can be easily handled in the computations.

Computer Programs

The primary purpose of this paper is to review selected computer programs in terms of uses for two purposes of a discriminant analysis, discrimination and classification. There exists today a variety of computer programs available to the educational researcher. There are a few very general multivariate programs (e.g., those by Elliot Cramer and by Jeremy Finn) that are available to users. One or more of a number of statistical computer "packages" are readily accessible at most institutions — BMD, OSIRIS, SAS, and SPSS are popular packages. IBM distributes a Scientific Subroutine Package (SSP) which includes a program designed to compute discriminant "functions." There are some books that list a number of computer programs (e.g., Veldman; 1967; Cooley and Lohnes, 1971; Overall and Klett, 1972). A book devoted exclusively to "discriminant analysis" by Eisenbeis and Avery (1972) also offers a set of computer programs. Individual computer programs are also available from writers: references are found in such journals as Educational and Psychological Measurement and Behavioral Science.

The discriminant analysis programs emphasized in this paper are those found in the widely used BMD package (Dixon, 1973); these are the 4M, 5M, and 7M programs. The titles given to these programs are: 4M, Discriminant Analysis for Two Groups; 5M, Discriminant Analysis for Several Groups; and 7M, Stepwise Discriminant Analysis. Because of the relationship between discriminant analysis in the two-group case and multiple regression analysis, the BMD 2R program, Stepwise Regression, will be included in the discussion. The three discriminant analysis programs will be reviewed individually, as well as relationships among these three and the regression program.

Discriminant Analysis for Two Groups

Beyond the basic computational results, the output from the 4M program includes the (unstandardized) discriminant function coefficients, a measure of distance -- Mahalanobis' D^2 -- between the two criterion groups (i.e., between the two group mean vectors, or centroids), the mean on the discriminant function for each group, and the discriminant function values for each individual (or case) in each group, printed in order of numerical value. [A value of an F-statistic, which is a transformation of the D^2 value, is also given which may be used to satisfy the third purpose of an analysis, separation, mentioned in an earlier section of this paper. In using this test one must assume p-variate normality.] A word about the discriminant function determined: computationally the coefficients are not found via the eigenanalysis so often associated with discriminant functions (Cooley and Lohnes, 1971, p. 246). However, the results are equivalent in the sense that the sets of coefficients obtained from the two analyses would be proportional.

If a purpose of the analysis is discrimination, as described earlier, little information is provided. If the D^2 value yields significance, then one may conclude there is one significant dimension of separation; this being represented by the determined discriminant function. No direct information is provided to 1) assess the contribution of each variable to the overall separation (which might be done by examining standardized coefficients), nor 2) aid in interpreting the discriminant function (where the variable versus discriminant function correlations might be used). With some arithmetic manipulation, however, this information may be obtained. To get the ith standardized coefficient

one can multiply the reported coefficient, a_i, by the (positive) square root of the ith diagonal element of the printed "SUM OF PRODUCTS OF DEV. FROM MEANS" (pxp) matrix; this matrix was denoted by W in the last section. The variable versus discriminant function correlations could be calculated from the information reported, but the computations would be fairly extensive — they involve matrix products. If one is merely interested in the ordering of the variables that would be determined by these correlations, a simple set of calculations need only be performed. It has been shown that this ordering is identical to that yielded by the ordering of the p univariate ANOVA F-values (Huberty, 1972). To determine the F-value for the ith variable the following expression is used:

$$\frac{d_i^2 \, (n_1 n_2)}{MS_i \, (n_1 + n_2)} \quad ,$$

where d_i is the difference of means on the ith variable, n_j (j=1,2) is the jth sample size, and MS_i is the error mean square for variable i. The d_i- and n_j- values are reported and MS_i may be found by dividing the ith diagonal element of W by $(n_1 + n_2 - 2)$.

Based on the output, the primary purpose behind the use of the 4M program is necessarily that of classification. Even then, the only classification that can be performed is that of the cases or individuals on whom the classification statistics were based. That is, there are no means of directly classifying "new" cases. Furthermore, the two sample covariance matrices are pooled in arriving at the classification statistic, which in this situation is merely the discriminant function. This implies that the population covariance matrices are assumed identical, which would make the use of the linear discriminant function quite appropriate. The

output does not provide sufficient information to determine whether or not this condition is met. The mean discriminant function value is reported for each group. Then, assuming equal costs of misclassification and equal prior probabilities of group membership, classification (of the cases already considered in determining the discriminant function) is simple. Cases whose discriminant function values are closest to the function mean of group j are assigned to group j. The sample proportion of misclassifications may be found by a mere count. If one is interested in obtaining an estimate of the true proportion of correct classifications he can use $\Phi(D/2)$, where Φ is the standard normal distribution function and D^2 is the reported Mahalanobis distance measure. [Here "function" is used in the mathematical sense.] This will yield an estimate that tends to be somewhat high.

The analysis yielded by the 4M program may be repeated using any number of specified subsets of the original predictor variables. If the user wants to discover what the results would be if one or more variables were deleted, the Selection Card is used.

Stepwise Regression

If the research situation is such that only two criterion groups are involved, as when the 4M program would be used, it might be well to consider the 2R program. The formal equivalence of two-group discriminant analysis and multiple regression analysis is well-known. That is, the regression coefficients obtained for the two-group situation, when the dependent variable is group membership, are proportional to the discriminant function coefficients. This statement holds when, for both analyses, the coefficients considered are those applicable to <u>raw</u> scores. With a

regression analysis, measures on the dependent variable are often taken to be 0 for all members of one group and 1 for all members of the other group. Coefficients (b_i's) comparable to the discriminant function coefficients are outputs of the 2R program; if care is taken in specifying the "F-level for inclusion," coefficients may be obtained for all p original variables. If desired, coefficients applicable to standardized scores (b_i^*'s) may be obtained by multiplying each reported coefficient by the product of $(n_1 + n_2)/\sqrt{n_1 n_2}$ and the reported standard deviation (s_i) of the variable in question; i.e., $b_i^* = b_i s_i (n_1 + n_2)/\sqrt{n_1 n_2}$. If $n_1 = n_2$ we have $b_i^* = 2 b_i s_i$.

Additional information may also be obtained from the 2R program (not using the zero regression intercept option) which may be useful in interpreting the separation between the two groups. First, regression equations consisting of different numbers of variables are determined in a stepwise manner. Thus an ordering of the variables in terms of their contribution to improved prediction is available. Subsets of variables may thus be selected, recognizing, of course, that a subset so selected may not be the best one of that particular size. Secondly, an ordering of the predictors according to discriminant function versus predictor correlations or, equivalently, to univariate F-ratios (or, in this case, absolute values of the univariate t-ratios) is possible. The F-value for the ith predictor is determined by

$$\frac{r_i^2}{1 - r_i^2} \quad (n_1 + n_2 - 2) \quad ,$$

where r_i is the point-biserial correlation between the ith predictor and the dependent (grouping) variable. The r_i -values are reported in the

(optional) output "CORRELATION MATRIX." If the composite versus ith predictor correlation coefficient is of interest, it may easily be obtained as

$$r_{i\,\frac{1}{R}} ,$$

where R is the multiple correlation coefficient based on all of the predictors (see Cooley and Lohnes, 1971, p. 55 or Mulaik, 1972, p. 404).

A measure of the distance between the two centroids may also be obtained from the 2R output. The value of D^2 is given by the relationship (see Porebski, 1966),

$$D^2 = \frac{(n_1 + n_2)(n_1 + n_2 - 2)}{n_1 n_2} \cdot \frac{R^2}{1-R^2} .$$

When the number of cases in each of the two groups is the same, output from the 2R program may also be useful for the purpose of classification. By assigning a 1 to cases in Group 1 and 0 to cases in Group 2 for scores on the dependent variable, classification results identical to those from the 4M program may be obtained by merely requesting the list of residuals to be printed. [The Subproblem Card must be set up so that all of the predictors are eventually included in the regression equation.] The proportion of correct classifications is found by counting the residuals closer to 1 for cases in Group 1 and residuals closer to 0 for cases in Group 2. As with the discriminant function values reported with the output from the 4M program, having the residuals from the 2R program enables the user to make interpretations regarding the misclassifications of particular cases.

Discriminant Analysis for Several Groups

The output from the 5M program consists of the basic statistics plus a generalized Mahalanobis D^2 value with an associated chi-square value, classification function coefficients and constants, posterior probabilities of group membership for each case, and a classification summary table. It should be noted that the generalized D^2 measure is not the same as the D^2 measure yielded by the 4M program; it is what Rao (1952, p. 257) denotes as his V-statistic. It turns out that in the two-group situation, $V=D^2 \cdot n_1 n_2/(n_1 + n_2)$. This statistic may be used as an alternative to Wilks' lambda statistic and, in the two-group situation, to Hotellings' T^2 statistic. It is appropriate at this point to discuss the resultant classification "functions." These are not the same as the usual discriminant functions (Cooley and Lohnes, 1971, p. 246). Rather, they are a modification of the "linear discriminant scores" discussed in Rao (1965, p. 488). The derivation of these functions is based on assumptions of multivariate normality and common covariance matrices. These functions do not take into account possible unequal prior probabilities of group membership, whereas Rao's do. In the two-group situation the differences of the corresponding coefficients obtained from the 5M program are proportional to the coefficients yielded by the 4M program (Rao, 1965, p. 489).

No information is printed which might aid the user in studying group separation. In the general g-group situation it is not possible to determine relative variable contribution nor dimensions of separation. It is not appropriate to rank-order the variables by examining the printed coefficients.

The 5M program is used basically for the purpose of classification. This classification analysis actually amounts to a "reclassification," in that each case is assigned to a population depending upon its function value which is based on the conglomerate of cases being assigned. That is, there are no means of classifying a "new" case into one of the predetermined categories. The classifications are determined by associated posterior probability (of group membership) values -- this is equivalent to basing the classifications on the largest function value obtained for each case. Potentially different prior probabilities of group membership are not considered.

Stepwise Discriminant Analysis

The last BMD program to be reviewed is 7M, Stepwise Discriminant Analysis. Of the discriminant analysis programs used in the reported literature the 7M program is probably referenced most often. Its widespread use might be attributed to the abundant amount of information yielded. Besides group means and standard deviations, within-groups covariance and correlation matrices are printed. At each step in the analysis various statistics are reported; a summary table is also printed, and plots of canonical -- actually linear discriminant function -- (deviation) scores are optional.

The "classification functions" computed in the 7M program are the same as those in the 5M program. [The constant terms yielded by the 7M program differ from the 5M constants, and are slightly in error.] It should be noted that the discriminant function coefficients based on the eigenanalysis of $W^{-1}B$ (assuming equal covariance matrices) are printed, along with the eigenvalues, following the summary table. In the printout

they are labeled "COEFFICIENTS FOR CANONICAL VARIABLE." [If p > g, only the first (g-1) sets of coefficients need be examined.] These coefficients may be scaled so that they are applicable to standardized scores by multiplying each coefficient by the (positive) square root of the product of (n-g) and the corresponding diagonal element of the printed "WITHIN GROUPS COVARIANCE MATRIX."

Considerable information is presented that may be used for the purpose of discrimination. First of all, the number of significant dimensions of separation may be determined by subjecting the reported eigenvalues to a significance test (see Tatsuoka, 1971, p. 165, and Harris, 1974). Following this the user can examine the plots of the discriminant scores to ascertain which groups are differentiated by which (significant) discriminant function. On the Group Label Card(s) different first letters for the g labels ought to be used. Only two-dimensional plots are given, but typically two functions account for almost all group separation. [It would aid in the interpretation of the functions if the variable-function correlations were available. No correlations are printed; however, correlations based on the total-group correlation matrix are obtainable thru the use of the 3D program, Correlation with Item Deletion. This would require the writing of a few FORTRAN statements to obtain the linear composites of the variables determined by the discriminant (not classification) function coefficients; this might be simpler than using transgeneration cards. These correlations may be used for interpretation as "structure coefficients" (Cooley and Lohnes, 1971, p. 248.] In addition to the scaled coefficients and the correlations, a third means of interpretation may be used. This is an assessment of variable contribution to group separation provided by the

ordering of variables entered into the analysis in a stepwise manner. [As might be expected, in the two-group situation the 2R and 7M programs yield identical orderings.] Further, the univariate F-statistics may be determined from the reported means and standard deviations (Gordon, 1973), or by using the means and the diagonal elements of the within-groups covariance matrix.

At each step statistics are reported which determine whether or not the variables entered significantly separate the criterion populations (in a mean vector sense). In addition, a matrix of F-values is given, each F-value being a transformation of a distance measure (Mahalanobis' D^2) between pairs of groups (Dixon, 1973, p. 241). The inverse of this transformation would yield distance measures which may be helpful in characterizing group differences. If, for example, distances between all pairs of g-1 of the groups are small, yet at the same time, the jth group is distinctly separated from the other g-1 groups, it is clear that the only differentiation taking place occurs between the jth group and its complement, i.e., the other g-1 groups.

As with many other discriminant analysis programs, including 4M and 5M, classification with the 7M program is usually carried out on the cases on which the classification statistics are based. Although results of classifying "new" cases would be more generalizable, results of the usual classification do provide descriptive information in that the total discriminatory power of the set of predictors may be assessed via the proportion of correct classifications. It is possible, however, with the 7M program to classify a group of cases which were not considered in determining the classification statistics. This is simply done by preceding that group size by a minus sign on the Sample-Size Card.

The classification procedure in the 7M program has the restriction of assuming equal covariance matrices (in that W is used) in determining the classification functions. However, it is different from the procedure in the 5M program in that it incorporates prior probabilities of group membership in computing posterior probabilities. The prior probabilities to be used may be specified on the Problem Card; the g priors most often used are given by the ratios of the group sizes to the total number of cases. Results of the classifications are given at each step in the analysis as well as after the final step.

Summary and Recommendations

Two purposes of a "discriminant analysis" are reviewed; those of discrimination and classification. The former pertains to a study of criterion group separation with respect to predictor variable contribution and dimensions of separation, while the latter involves the assignment of cases (individuals or objects) to criterion populations. The usual requisite conditions of normality, homogeneity of dispersion, and equal costs of misclassification are discussed. The primary purpose of this paper was that of reviewing a set of computer programs designed to carry out a "discriminant analysis" in light of purposes and requisite conditions. Interpretation of the outputs from these programs is covered, along with similarities and differences across program outputs.

When using the BMD discriminant analysis programs it is recommended that multiple analyses be made; reanalyzing data with the same program, and, when appropriate, with different programs. The programs may be used more than once by varying some of the options available; for example, using different variable subsets in the 4M program; using different

F-levels, or variable selection criteria in the 7M program. Running analyses on a given set of data using different BMD programs is also helpful; for example, obtaining outputs from 2R, 4M, and 7M on the same data. It should be noted that such multiple runs, using the same or different programs, on the same data may be accomplished by a single submission of the data to a computer center.

Three further recommendations may be made when using the BMD programs. One is to use appropriate prior probabilities in the 7M programs. Unless results from past research on similar variables is available, and unless other theoretical considerations can be used to assess prior probabilities of group membership, it is well to use priors of $n_j/\Sigma n_j$. Another recommendation pertains to estimation of proportions of correct classifications or of misclassifications. If the number of cases to be classified is large enough, it would be well to use the validation procedure afforded by the 7M program to classify new cases (see, however, Horst, 1966, pp. 139-140). To do this one can use what is called a "hold out sample." A third recommendation is to examine multi-univariate analyses to screen data prior to using, say, the 7M program (see Huberty, 1974).

The BMD programs yield information which may be used in subsequent calculations to determine statistics for more complete interpretation. For example, discriminant coefficients applicable to standardized scores may be determined from output of both the 4M and 7M programs, as well as from the 2R output in the two-group situation. Univariate F-values are also obtainable from the 4M and 7M output, as are correlations between predictors and discriminant functions. These three statistics, plus the ordering of variables entered as determined by the 2R and 7M programs, can be examined in assessing variable contribution to

to separation and in interpreting the discriminant functions (see Huberty, 1971; Tatsuoka, 1973, p. 280).

Depending upon the purpose of a study and resources available the researcher might do well to use other computer programs in lieu of, or in addition to, the BMD program(s) selected. In this way other statistics may be examined, e.g., test statistics and classification statistics. In particular, it is advised that programs using quadratic classification functions which do not require equal covariance matrices be selected when the data are such that linear functions are inappropriate. It is of interest to note that a new BMD program is now available; this program requires some special hardware, and may be obtained for a small cost. This new program, which is discussed by Dixon and Jenrich (1973), has three very promising added features: provision for (1) more meaningful graphic interpretation of results, (2) the handling of the unequal covariance structure problem, and (3) specifying relative costs of misclassification as well as prior probabilities for each group.

FOOTNOTE

[1] It ought to be noted that this use of the term "function" is not mathematically correct. However, tradition will be followed in this paper by using the term to mean a linear composite.

REFERENCES

Bausell, R. B. and Magoon, J. Expected grade in a course, grade point average, and student ratings of the course and the instructor. Educational and Psychological Measurement, 1972, 32, 1023-1023.

Chastian, K. Prediction of success in audio-lingual and cognitive classes. Language Learning. 1969, 19, 27-39.

Cooley, W. W. and Lohnes, P. R. Multivariate data analysis. New York: Wiley, 1971.

Dixon, W. J. (Ed.) Biomedical computer programs. Los Angeles: University of California Press, 1973.

Dixon, W. J. and Jenrich, R. I. Computer graphical analysis and discrimination. In T. Cacoullos (Ed.), Discriminant analysis and applications. New York: Academic Press, 1973. Pp. 161-172.

Eisenbeis, R. A. and Avery, R. B. Discriminant analysis and classification procedures. Lexington, Mass.: Heath, 1972.

Fisher, R. A. The use of multiple measurements in taxonomic problems. Annals of Eugenics, 1936, 7, 179-188.

Goldman, R. D. and Warren, R. Discriminant analysis of study strategies connected with college grade success in different major fields. Journal of Educational Measurement, 1973, 10, 39-47.

Gordon, L. V. One-way analysis of variance using means and standard deviations. Educational and Psychological Measurement, 1973, 33, 815-816.

Harris, R. J. Multivariate analysis of variance. Paper presented at the annual meeting of the American Educational Research Association, Chicago, April, 1974.

Horst, P. An overview of the essentials of multivariate analysis methods. In R. B. Cattell (Ed.), Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally, 1966. Pp. 129-152.

Huberty, C. J On the variable selection problem in multiple group discriminant analysis. Paper presented at the Annual Meeting of the American Educational Research Association, New York, February, 1971.

Huberty, C. J Regression analysis and 2-group discriminant analysis. Journal of Experimental Education, 1972, 41, 39-41.

Huberty, C. J Discriminant analysis. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, April, 1974.

Keenen, C. B. and Holmes, J. E. Predicting graduation withdrawal and failure in college by multiple discriminant analysis. Journal of Educational Measurement, 1970, 7, 91-95.

Kendall, M. G. Discrimination and classification, In P. R. Krishnaiah (Ed.), Multivariate analysis. New York: Academic Press, 1966. Pp. 165-185.

Melton, R. S. Some remarks on failure to meet assumptions in discriminant analysis. Psychometrika, 1963, 28, 49-53.

Mulaik, S. A. The foundations of factor analysis. New York: McGraw-Hill, 1972.

Nicholson, E. Predictors of graduation from college. ACT Research Report No. 56. Iowa City, Iowa: The American College Testing Program, 1973.

Overall, J. E. and Klett, C. J. Applied multivariate analysis. New York: McGraw-Hill, 1972.

Porebski, O. R. Discriminatory and canonical analysis of technical college data. British Journal of Mathematical and Statistical Psychology, 1966, 19, 215-236.

Rao, C. R. Advanced statistical methods in biometric research. New York: Wiley, 1952.

Rao, C. R. Linear statistical inference and its applications. New York: Wiley, 1965.

Rock, D. A., Baird, L. L., and Linn, R. L. Interaction between college effects and students' aptitudes. American Educational Research Journal, 1972, 9, 149-161.

Stahmann, R.F. Predicting graduation major field from freshman entrance data. Journal of Counseling Psychology, 1969, 16, 109-113.

Tatsuoka, M. M. Multivariate analysis. New York: Wiley, 1971.

Tatsuoka, M. M. Multivariate analysis in educational research. In F. N. Kerlinger (Ed.), Review of research in education. Itasca, Ill.: Peacock, 1973. Pp. 273-319.

Veldman, D. J. FORTRAN programming for the behavioral sciences. New York: Holt, Rinehart and Winston, 1967.

Whellams, F. S. Musical abilities and sex differences in the analysis of aural-musical capacities. Journal of Research in Music Education, 1973, 21, 30-39.